E-599E
c. 2.

Measuring the Quality of Data in the 1981 Census of Population and Housing of Canada

Statistics Canada
December 1983

(Prepared by R. Burgess, Census and Household Survey Methods Division)

Measuring the Quality of Data in the 1981 Census of Population and Housing of Canada

1. Introduction

The Census of Canada, given its size and significance to the work of Statistics Canada, and to data users, presents special problems and responsibilities to ensure that the data can be properly analysed and interpreted. This, of course, requires that the user be provided with the basic definitions of variables and relevant terminology. These, however, frequently will not be sufficient.

The interpretation or use of the data is not bound to a given Census. For example, changes in population characteristics and distributions from earlier Censuses are of fundamental interest. Such comparisons are meaningful only to the degree that the Census data are historically comparable. The comparisons may be affected by changes in questions, concepts or procedures.

The use of the data may also be affected by error. The inevitability of some non-response, and of coverage, response, processing and other errors in the data, is commonly accepted. The geographic size of Canada and the mobility, and varied social and cultural characteristics of its population, present potential for error which requires special consideration - a consideration which must be viewed in light of increasing use of data for smaller geographic areas and for sub-groups of the population.

To encourage the most effective use of the data, the output of the Census should include information on the concepts and methods which historically underly the data and information on data quality.

The purpose of this paper is to describe the evaluation programme undertaken for the 1981 Census of Canada. The objective of this programme was to provide users with the fundamental information on comparability and quality.

2. 1981 Census Evaluation Strategy

The evaluation of the 1981 Census data was planned and carried out in a manner consistent with the practices of previous Censuses, while at the same time extending and reformulating these practices to better achieve the objective of providing better service to the user. To this end planning was done using five basic guiding principles:

i) counts and data must be analysed and reported on, from the perspective of quality and historical comparability, prior to release;

ii) anomalies and conceptual or significant procedural changes should be identified at the time of release and documented as part of the disseminated data;

iii) where feasible and practical, estimates of error for potentially significant sources of error should be produced;

iv) estimates of error produced must be meaningful to the user and provide him/her with guidance on the potential impact of error on the use of specific data;

and v) an assessment of the quality of the data and of the impact of the methodology and procedures of the Census on the quality and comparability of the data must be documented, published and disseminated as part of the regular Census publication programme.

For the first two guidelines there was a timeliness problem to overcome. An evaluation of the data based upon statistical studies which produce estimates of error is a time consuming process. To withhold dissemination of the Census data until such measures were available would not have been acceptable to users and would have seriously depreciated the relevance and credibility of the data as a current indicator of the characteristics and distributions of the population.

For the third and fourth guidelines there were strict limitations on the extent of evaluation that could be conducted from the perspective of cost, timeliness and feasibility. There were similar constraints on the level of geographic detail and population sub-groupings for which consideration for evaluation could reasonably be given.

To overcome or to mitigate these constraints two approaches to the evaluation were undertaken - the one termed certification, the other data quality measurement. The certification approach was directed towards meeting the first two guidelines, i.e., towards establishing, prior to release and dissemination of data, a basic knowledge of the quality of the data, and of the impact of procedural or conceptual changes on historical comparability. As a consequence of providing at least an initial view of the data quality, the certification process identified specific problem data areas where less timely but more detailed and more rigorous evaluation could be conducted under the data quality measurement process.

Data quality measurement was directed towards meeting the third and fourth guidelines, i.e., towards providing basic measures of data quality, and providing quantitative data on the impact of Census operations on the data. This included identifying sub-populations with significant error or potential error and providing measures of this error, and conducting detailed evaluations of problem areas identified during the certification process.

The last guideline was met by including in the formal Census products and services programme publications devoted to documentation and discussion of the quality of the data.

3. Certification

The objective of the certification process was to identify data quality problems which required rectification of the data before dissemination, notification to users at the time of dissemination, further evaluation and/or some

combination of these. This was achieved by way of a macro assessment, for major geographic levels, of universe counts (i.e., counts of persons, households, families, etc.) or proportions and of univariate and multivariate distributions, and of measures derived from these distributions. Such assessments were facilitated primarily by making comparisons with alternative sources of data and with data from previous Censuses, either through direct comparison of distributions or through the use of projections especially prepared for this purpose.

3.1 Scope of Certification

Given that there were more than one hundred universes and variables to be certified and that data for these would be available for many different levels of geographic detail, the timeliness requirements dictated that only major aggregations of the data be checked. For most variables certification was carried out at the Canada and province level and only sub-provincially if higher level checks indicated a potential data quality problem, if by the nature of the variable provincial checks were not meaningful (e.g., Place of Work), if the variable was considered to be of special importance or if there was a legal requirement for sub-provincial certification. In the latter category were included the population counts (verified to the Census Tract level - 3,000 to 10,000 persons) and the Mother Tongue variable verified to the Federal Electoral District (an area represented by a member of Parliament).

Univariate distributions for all variables available to users through published tables or by special request were certified. Multivariate distributions for variables frequently associated, or for which there was some predictable correlation, were also certified.

3.2 Method of Certification

Certification was carried out by groups of subject matter experts, each group dealing with one or more of sixteen sets of related universes and/or variables. Prior to the availability of the Census data to these groups, alternative sources of data and a strategy for verification and acceptance or rejection were established.

A variety of alternative data sources were used. Sources other than earlier Censuses included data from Statistics Canada surveys, administrative or survey data from federal, provincial and municipal government departments and agencies, and data from a variety of publications of a specialized or technical nature.

Typically, as an alternative source, the 1971 and 1976 Censuses were used. In the case of housing and household variables, data back to the 1961 Census were used. For family data reference was in some cases made back to the 1941 Census.

Two general verification strategies were used. The one was a comparison between the 1981 Census data and that of the alternative source. Tolerance levels were pre-set for the degree of acceptable deviation of distributions from comparable distributions for alternative sources of data. These tolerance levels were set with consideration for comparability of concepts and effective dates, and the limitations of the universe, methodology, data

quality, etc., of the alternative source. As an example, the certification of age/sex/marital status variables used independent demographic population estimates as one alternative source. The tolerance levels for the Census data were set at ± 5% of the comparable population estimate for all cells of size 100,000 or more: and ±10% for cells in the range of sizes 10,000 to 100,000. Smaller cells were examined on an ad hoc basis.

When a cell or other measure exceeded tolerance it was necessary for a rationale to be provided to explain the deviation. In some cases such an explanation was readily found in other Census data distributions (e.g., because of unexpectedly high interprovincial migration certain age-sex cells within a particular province were larger than projected). In other cases subsequent investigations showed the deviations to be due to an error in the Census data; in others due to an error in the data from the alternative source.

The second verification strategy was applied when there was no previous Census precedent to project from, or if projection from a previous Census was not entirely meaningful, and if external alternative sources were not adequately comparable or were not available. For such variables it was necessary to provide a basic rationale for the distributions. This interpretative verification was based upon logical criteria being placed on the data. For example, for one first time Census question, related to dwelling ownership, one particular response would not be expected to be common except for certain dwelling types in urban areas. Significant occurrence of this response in other areas, or for other dwelling types, would have indicated a probable error.

The data comparison verification strategy accounted for more than 80% of the certification. The interpretative technique was used for the remainder.

As part of the certification process, for all variables comparisons were made between distributions for edited imputed data and for unedited unimputed data. Thus to a limited extent some separate assessment of Census processing and procedures was made.

4. Data Quality Measurement

The data quality measurement process was an examination of the error and/or change introduced by five sources of potential error in the Census -- coverage, response, processing (other than edit and imputation), edit and imputation, and sampling and weighting. These sources were intended to encompass all aspects of data collection and processing of the Census data.

For each of the sources, studies were undertaken to provide either estimates of error or to quantitatively describe the impact of operations or processes on the data. A decision on what studies to conduct, among the viable options, was based upon an assessment of the significance to the uses of Census data of the potential error being examined and of the cost effectiveness of the option.

4.1 Nature of Operations

The measures of data quality were produced by way of four general types of study:

i) Enumeration checks: enumeration checks were statistical studies, on a sample basis, of some particular aspect of Census enumeration and involved some degree of re-enumeration. Such checks were carried out primarily as part of coverage studies.

ii) Parallel operations: parallel operations were studies which were performed, independently and using a sample of Census data, in parallel with a Census operation. A comparison of the Census output with that of the sample operation permitted estimation of the error introduced. This type of study was carried out to evaluate processing.

iii) Alternative data source comparisons: comparisons with alternative data sources involved micro-matches, on a sample basis, of Census records with those of an alternative source. This type of study was carried out to evaluate response and edit and imputation.

iv) Analysis of Census data: analysis of the Census data involved some form of calculation using finalized and/or intermediate (unedited) Census data. Such studies were carried out to produce response rates and estimates of sampling error and to evaluate edit and imputation.

4.2 Coverage Component

The most fundamental purpose of the Census is to provide counts of the population for various geographic levels. Errors in coverage therefore, have special significance and a comparatively large proportion of evaluation resources was applied to estimation of these errors.

A coverage error occurs whenever a person, a household or dwelling is missed completely (undercoverage), counted more than once (overcoverage), or whenever a unit not in a Census universe is counted (overcoverage). Generally overcoverage is expected to be fairly rare in relation to undercoverage and is more difficult to evaluate. The coverage component of data quality measurement programme therefore was directed primarily, but not exclusively, to estimation of undercoverage error.

An assessment of four elements of coverage error was made:

i) population and private household undercoverage, which were evaluated by the Reverse Record Check;

ii) population undercoverage of persons temporarily absent from their usual place of residence, which was evaluated by the Temporary Resident study;

iii) population and household undercoverage of persons whose usual place of residence was misclassified as unoccupied, which were evaluated by the Vacancy Check;

and iv) dwelling overcoverage due to structures not in the housing stock (because they were under construction or renovation, seasonal dwellings, etc.) being misclassified as unoccupied

4.2.1 The Reverse Record Check

The Reverse Record Check (RRC) has been conducted as part of the Census of Canada since 1966. The specific objectives of the RRC were i) to provide estimates of population undercoverage at the province level; and ii) to provide estimates of private household and population undercoverage for various sub-groups of the population and housing universes.

There were two key elements of the RRC methodology: i) the construction of an independent list of all persons who, it was believed, should have been enumerated in the census; and ii) the tracing/follow-up operations which were undertaken to identify intercensal deaths and emigrants on the constructed list, and to ensure that the Census Day usual place of residence for persons who should have been enumerated was known.

Conceptually these two elements permitted matching to the Census to determine whether or not the persons who should have been were enumerated. The construction of the list and subsequent tracing, follow-up and matching were carried out for a sample of persons.

The independent list had four components or frames:

i) all persons enumerated in the previous (1976) Census;

ii) all intercensal births - obtained from vital statistics records;

iii) all intercensal immigrants - obtained from immigration records;

and iv) all persons missed in the previous (1976) Census - obtained from the 1976 RRC.

A sample of persons was selected from all frames except for the fourth, as this was already represented by the sample of cases from the 1976 RRC which were classified as missed in the 1976 Census. A total sample of approximately 36,500 persons was selected.

After sample selection extensive operations were conducted to trace and to follow-up on the selected persons. For most selected persons an available address (e.g., from the previous Census) was first searched in the Census documents. If the selected person was not found enumerated at this address, then contact was attempted through telephone interviews and through searches of city directories and of various government administrative records. Census Day addresses obtained for selected persons through contact were then searched. If a selected person was not found enumerated at the given address a follow-up interview was conducted to obtain information on possible alternative addresses and to obtain certain Census data for him/her and the household. This information was used to confirm that the selected person was missed in the Census and, if he/she was, as input to estimates of undercoverage.

The sample cases were classified, as a result of these operations, into one of five basic results: i) enumerated in the Census, ii) missed in the Census, iii) deceased before Census Day, iv) emigrated before Census Day or living abroad on Census Day, and v) selected person not traced. The "not traced" category was in effect "non-response" and represented those cases sent

to a tracing operation but not located. Among all selected persons 3.4% were classified to this category. These cases were removed from subsequent estimation by way of weight adjustment procedures.

4.2.2 The Temporary Resident Study

The Census of Canada is "de jure" and not "de facto". Thus persons temporarily absent from their usual place of residence are still to be enumerated at their usual place of residence. All persons in Canada on Census Day who were in the Census universe (i.e., were not foreign residents) and who were staying temporarily at some residence (e.g., hotel, hospital, or jail) were enumerated there on a special Census form called a Form 3.

After collection, a 2% sample of completed Forms 3 (approximately 10,000) was selected and tracing undertaken to identify among the regular Census questionnaires the usual place of residence of the selected person. This was facilitated by address information requested on the Form 3. It was then determined, using data collected on name, age, sex and marital status, whether the selected person had been enumerated at his/her usual place of residence.

Estimates of the number of persons missed by age, sex and marital status groups for broad geographic areas within provinces were produced. These estimates were used to adjust or to "correct" the Census population counts for this particular incidence of undercoverage. This adjustment was effected through a weight adjustment process using randomly selected donor records with given characteristics (age, sex, marital status) within given geographic areas. By this process the weight of the donor record was incremented by one to thus represent both itself and the missed person. Nearly 118,000 persons were added to the initial population counts through this methodology. This was approximately 0.5% of the final population count and 23% of all persons enumerated on a Form 3 at a place of temporary residence.

It was because persons enumerated as temporary residents represented an easily identifiable group known to be subject to a high rate of undercoverage that this study and subsequent adjustment were undertaken.

4.2.3 The Vacancy Check

Dwellings enumerated as unoccupied formed another group within the Census which was known to be subject to a relatively high rate of coverage error. There were in this case two aspects of coverage error to be examined. First, dwellings enumerated as unoccupied may have been in fact occupied on Census Day. This would represent undercoverage of persons, households and occupied dwellings, and overcoverage for unoccupied dwellings. Such misclassification may have resulted because the household moved close to Census Day or was absent for some period of time (e.g., on vacation, or staying at a summer residence). Second, structures enumerated as unoccupied dwellings may not have been suitable for year-round occupancy on Census Day and thus should not have been classified as dwellings. This would represent overcoverage for unoccupied dwellings and for total dwelling stock.

In order to estimate these elements of coverage error a sample re-enumeration of dwellings or structures enumerated as unoccupied was undertaken soon after completion of the Census collection operation. The work of approximately 1250 enumerators involving approximately 14,000 "unoccupieds", was

selected for this study.

The re-enumeration was carried out through intensive interviewing of current occupants (if the dwelling was currently occupied), building superintendents, and neighbours. From the results of the interviews it was established whether the structure was actually occupied on Census day. In the event that the structure was occupied, it was established, through searching of addresses provided, whether the occupants were enumerated elsewhere. The number of persons in the household and the dwelling type (e.g., apartment) were also determined.

In the event that the structure was unoccupied on Census day it was established whether the structure was a dwelling. To be a dwelling a structure had to be suitable for year round occupancy as of Census Day. This meant that it had to include basic facilities. For example, a house under construction or renovation, and an unheated summer cottage would not be considered dwellings.

Two primary sets of estimates of coverage error were the results of the study. First, estimates of the number of households missed by household size, by dwelling type, by urban-rural areas within provinces were produced. (The estimate of missed households represented also a corresponding amount of overcoverage of unoccupied dwellings). These estimates were used to adjust or to "correct" the Census population, occupied dwelling (household) and unoccupied dwelling counts for this particular incidence of under and overcoverage. The adjustment of population and occupied dwelling counts was effected through a weight adjustment process similar, in general terms, to that used for the temporary resident adjustment. In this case donor records were household records selected on the basis of household size and dwelling type within given geographic areas. A corresponding number of unoccupied dwellings were deleted from the data base. Approximately 44,500 households and 82,000 persons were added to the initial population and household counts through this methodology. This was approximately 0.3% of the final population count and represented a 10% misclassification rate for structures enumerated as unoccupied.

Second, various estimates of the number of structures misclassified as dwellings were produced. These were not used for adjustment but were used in a Census data publication to caution the user on overcoverage in total dwelling stock counts and in unoccupied dwelling counts. On an after adjustment basis the overcoverage rate for unoccupied dwellings was estimated to be nearly 20%. Approximately one half of this was due to the erroneous enumeration of dwellings under construction or renovation.

4.2.4 Net Undercoverage Error

The results of the RRC matching to the Census reflected only the completeness of the Census counts prior to the adjustments for temporary residents and misclassified unoccupied dwellings. The RRC estimates thus had to be adjusted to account for these statistical additions to the data base. This yielded a net 2% rate of population undercoverage for the 1981 Census. A level comparable to that of the three previous Censuses.

4.3 Response Component

Three general studies were conducted to evaluate response quality - a

response rate study, a content study based upon the RRC sample, and various problem area investigations.

4.3.1 Response Rate Study

Response rates at the Canada level for each variable were produced. These response rates were only general indicators of the level of quality but provided a measure of the extent to which the Census data were based upon responses from respondents and to some extent of the amount of imputation required.

For a sample of Census records, more extensive univariate and multivariate response rates were calculated. These were produced to identify groups of the population with high rates of non-response. High rates for particular groups gave indication of the potential for bias in data representing such groups. Correspondingly, for the majority of groups, low rates of non-response gave indication of low potential for non-response bias.

4.3.2 Reverse Record Check Content

The RRC content study used all selected person records which were classified in the RRC as "enumerated" in 1981. For these records comparisons between the responses for the person in the 1981 Census and those from the 1976 Census (or from administrative records for persons not enumerated in the 1976 Census) were made. Comparisons were made for those variables which should be stable over time (e.g., date of birth, sex, mother tongue) and for those for which change could only be in one direction (e.g., highest level of schooling attained). The primary measures produced were estimates of relative bias and gross difference rates.

4.3.3 Problem Area Analysis

Studies were conducted as a result of problems being identified during the certification process. Results were typically an explicit identification of the problem but without necessarily being able to provide estimates of the precise degree of error.

4.4 Processing Component

Two studies were conducted to evaluate the processes of coding, data capture and manual editing.

4.4.1 Coding Bias Study

Coding was necessary to transform various write-in entries (e.g., industry, occupation) into numeric codes which could then be captured. A study of the bias introduced by coders was conducted.

The estimation of coding bias was facilitated by comparing codes assigned by coders with those which would have been assigned by "expert" coders. This was done on a record by record basis for individual variables and involved a sample re-coding operation. In producing estimates it was assumed that the expert was correct. The reliability of the results would be limited to some unknown extent by this assumption, by the fact that only a sample of person records could be used in the study, and by the infrequency of occurrence of

write-ins for some variables and of individual codes.

4.4.2 Data Capture Evaluation

The Census data were captured by a key entry process. This method of capture methodology had not been used in the previous Census, thus it was important to show that this process was not a significant source of error.

For this study a sample of Census questionnaires amounting to data for approximately 33,000 persons was recaptured independently of the main flow of Census data. The re-keyed data were then processed through the remainder of the capture system which included basic manual consistency and head count edits. Changes to the data for the sample, as a result of this processing, were made by experts.

Once re-processed the sample data records were matched to the corresponding records captured as part of the main Census process. Any discrepancies between the two versions of the data were outputted for review by the expert staff. Discrepancies due to an error in the main Census process were flagged. Estimates of processing error were then produced, by variable, using the flags as indicators of error.

4.5 Edit and Imputation Component

Three general studies were conducted to evaluate the impact of edit and imputation on the Census data. These studies parallelled those undertaken as part of response evaluation.

4.5.1 Data Relationships

The Data Relationships study provided quantitative measures of the extent of imputation - both for edit failures and for non-response. While such measures were not direct estimates of error they provided a strong indication of the quality of the data. Since an edit and imputation process is directed to detectable error (including non-response), given any fixed set of processing procedures, the smaller the proportion of change or of imputation the more reliable the data should be. To some extent this must, of course, lead to some assessment of the actual rules of the process.

Estimates of the rate of imputation were produced for each variable according to each final value. Similar estimates on a multivariate basis using combinations of imputed and unimputed values were produced. These were used to identify groups of the population for which a relatively high rate of imputation occurred and to determine the propensity for imputing particular values. This in part constituted an assessment of the edit and imputation rules applied.

Some direction to this process was given by the certification activity which identified particular areas for which edit and imputation had a significant impact on the data and by the results of the work on response rates.

4.5.2 Reverse Record Check Match

Once the matching process for the RRC Content study had been done it was

possible to use these same matched records to assess the impact of edit and imputation. Essentially this was identical to the RRC Content methodology except instead of matching initial Census responses, the RRC responses (either from the previous Census or administrative records, or those collected as part of the RRC process) were matched to imputed values where imputation occurred.

Separate estimates were produced for imputation required to resolve edit conflicts and that required for non-response. Estimates produced for the edit conflict cases were assessed in conjunction with corresponding RRC Content results.

4.5.3 Problem Area Analysis

The process for investigation of specific problems in edit and imputation was similar to that for suspected response problems. Typically the problem was identified during the certification or through another study and resulted in a detailed examination of the specific edit and imputation procedure in conjunction with circumstances of response and collection.

4.6 Sampling and Weighting Component

While basic data for the 1981 Census were to be collected for 100% of the population an extensive amount of data was to be collected only from a 1 in 5 random sample of households. There was thus a need to evaluate the impact of sampling, and the consequent weighting process, on the quality of data. Three primary studies were conducted for this evaluation.

4.6.1 Sampling Bias Check

The requirements of the sample design for the Census, and for corresponding sample selection procedures to be followed by the enumerator, should have ensured that an unbiased sample was selected. The Sampling Bias Check was conducted to evaluate the extent to which sample selection was implemented according to the intended design.

This evaluation was made by comparing for all Enumeration Areas (an EA was a geographic area which was the collection assignment for one enumerator) the distribution of household characteristics of sampled households with those for all households in the EA (using data collected for all households). Statistical tests were applied to these comparisons in order to determine if differences, among all EAs in given geographic areas, deviated significantly from what could be expected for a simple random sample of households.

4.6.2 Sample/Population Consistency Check

The Census weighting procedure was carried out for groups of contiguous EAs called Weighting Areas (WAs). A WA typically would have had a population count in the range of 3000 - 7000 persons. The weighting methodology made use of the various population and housing data collected for all households by employing a raking ratio estimation procedure. This procedure was what might be considered a two dimensional ratio estimation procedure with each dimension being implemented alternately and iteratively. For characteristics used in this adjustment process, estimates based upon the sample data, at the WA level or higher, would closely approximate the corresponding counts that would be obtained from the corresponding complete or 100% population of persons or

households. For data collected on a 100% basis, therefore, this procedure should have yielded considerable consistency between tabulations based upon the 100% data and identical tabulations based upon sample data.

The Sample/Population Consistency Check was conducted to evaluate the extent of this consistency and in particular to determine whether consistency at the WA level had been achieved at the expense of consistency at sub-weighting area levels. This evaluation was made by comparing, for various geographic levels down to the EA level, the 100% counts for a variety of 100% variables to the corresponding sample data estimates. As in the Sampling Bias study, statistical tests were applied to these comparisons.

4.6.3 Sampling Error

The evaluation of error introduced by sampling in the Census was a fundamental undertaking. For most Census variables, estimates of variance (incorporating sampling variance and some components of response and processing variance) were produced. The estimates were based upon a ten percent sample of WAs, and were produced using a Taylor Series expansion estimator which incorporated the effect of the weighting procedure. Separate sets of estimates were produced by province.

5. Publication of Data Quality Information

There were several ways in which information on the quality of data were (or are to be) disseminated. All data disseminated as part of the planned Census products programme were accompanied by standardized textual material on the quality of data. This included a discussion of the potential sources of error and cautioned the user to be aware that the data are subject to error, particularly in the case of smaller cell values. In some cases Canada level estimates of error were included (e.g., the estimated rates of population and private household undercoverage and response rates).

The textual material to accompany data collected on a sample basis included a discussion of the potential error due to sampling and a table of sampling errors (by cell size) with a statement on the use of these measures. If, in addition, the data being disseminated included average income cells, then the data tables also included corresponding estimates of standard error of the average income estimates.

This information was, of course, in addition to the various types of notes discussed earlier, and the basic definitions of variables and terminology. The material of this nature, however, varied according to the data included in the particular tabulation package.

For data disseminated as a result of a request by a user, basic textual material, including cautionary notes, was sent with the requested tabulations. In addition, at the time of request, the quality of the data requested and matters of comparability, definition, etc., were discussed thoroughly with the user.

Any textual material sent with Census data tabulations and any other relevant publications referred the user to two publications devoted to documentation and discussion of the quality of the Census data. These two

publications represent the most comprehensive form in which the quality of the data its historical comparability, and the impact of '81 Census methodology on the quality and comparability are discussed. One publication deals with the data collected from the total population; the other deals with the data collected from the sample of households.

These publications briefly describe those aspects of the Census methodology which impact on the quality and comparability of the data and present the results of the data quality measurement process by source of error. The largest portion of these documents is, however, a universe by universe and variable by variable assessment of the Census data. This assessment is based upon an analysis of certification results, re-considered in light of the results of the data quality measurement process.

6. Effectiveness of the Evaluation Programme

Planning and developmental work for the evaluation of the 1986 Census are going on at this time. One basis for this work is an assessment of the effectiveness of the 1981 programme.

6.1 Certification

For 1981 the certification process attempted to evaluate the historical continuity and current consistency of the Census data. It was directed towards detecting significant bias in the Census data. Such bias would be detected providing certain conditions were met:

i) there was not a consistent historical bias or bias in common with alternative sources;

ii) concepts, methodologies, etc. of alternative sources were well understood;

iii) projections used for comparisons were well formulated and valid;

iv) tolerance limits reasonably represented non-bias deviations from the alternative source;

and v) the bias was relatively large so as not to be obscured by variances associated with the Census data and the data of the alternative source.

Assurances that such conditions were met for any given variable either had to be provided by reasonable evidence from the certification process itself that the data for the particular variable did contain a significant bias, or by the results of the data quality measurement studies. The certification approach was effective for variables which historically have been well studied, and for variables which are stable over time or for which change is relatively easy to predict (e.g., age, sex and housing variables). In general, the data quality measurement results for these types of variables supported the certification conclusions. There were two important cases for which the certification results were not supported. These were determined to be cases for which there was a historical bias (response or enumerator) in the Census

data which invalidated the certification comparisons.

The approach was less effective for cultural and other variables which may have been subject to significant response errors, but which were difficult to evaluate. Generally, for such variables there was a lack of sound or comparable alternative data sources, or change over time was difficult to predict. There was often a lack of meaningful data quality measurement of response error for such variables. For these cases, the conclusions of certification may remain subject to some debate.

The certification process, therefore, did leave some potential for data to be released and to contain important, but unknown (and unsuspected) error. This potential, however, had been substantially diminished by the certification. At the same time, the analytical work which was part of certification, provided, prior to release, a knowledge of what phenomena (real or otherwise) the Census data were describing. This directly enhanced the effectiveness of communication with users.

6.2 Data Quality Measurement

For most, if not all, aspects of the Census which were evaluated the data quality measurement studies provided meaningful and useful results. The limitations of this part of the programme were related to its breadth, or detail of application. Three limitations are worth noting. First, there are potentially important sources of error for which there are no estimates. For example, for many variables there are no estimates of response bias, and there are no estimates of population overcoverage or of the impact of illegal aliens on the quality of counts.

Second, estimates of error are not generally applicable for small area data. While some estimates (e.g., estimates of sampling error) may seem to be applicable for any cell size, the actual error may be greatly affected by localized or specialized circumstances. In part this is a problem of the sampling error associated with the estimates of data quality.

Third, and to some extent indistinguishable from the second, there are few estimates or measures of total or aggregate error. This will not necessarily be a serious problem for large cells values, if all major potential sources of error can be evaluated (regardless of the level of error). It will become an increasingly serious problem as cell sizes decrease, and in particular for small area data.

7. Conclusion

The evaluation programme for the 1981 Census included important improvements on similar programmes for earlier Censuses. There was some increase in the scope and detail of evaluation. The most significant improvements were, however, in the application of a formal certification process, and in the scope, nature and packaging of information for users.

It would be obvious to conclude that the limitations of the 1981 evaluation programme should be addressed for the 1986 Census. It must be remembered, however, that improving data quality evaluation is an evolutionary process functioning within an environment of revolutionay changes in data

usage and manipulation. It is unlikely that the limitations of the 1981 programme can be overcome in the near future, while demand for data will increase. It is therefore essential that there be a continuing concern for data quality in the design of Census questionnaires, methodology and operations. The quality of the Census data will thus be maximized even where this quality cannot be measured precisely.

(4)
E-599F
C.2.

Mesure de la qualité des données du recensement de 1981 visant la population et le logement au Canada

Statistique Canada.
Décembre 1983.

Préparé par R. Burgess, Division des méthodes de recensement et d'enquêtes-ménages.)

Mesure de la qualité des données du recensement de 1981 visant la population et le logement au Canada.

1. Introduction

En raison de son ampleur et de l'importance qu'il revêt pour les travaux de Statistique Canada et pour les utilisateurs des données censitaires, le recensement du Canada pose des problèmes particuliers et oblige le Bureau à s'assurer que les données obtenues peuvent être correctement analysées et interprétées. À cet égard, il faut nécessairement fournir à l'utilisateur les définitions fondamentales des variables, ainsi que la terminologie pertinente. Néanmoins, il arrive souvent que cette documentation soit insuffisante.

L'interprétation ou l'utilisation des données censitaires ne se limite pas à un recensement donné. Par exemple, les changements qui se produisent dans les caractéristiques démographiques et la répartition de la population par rapport aux recensements antérieurs sont d'un intérêt primordial . Ce genre de comparaisons n'est toutefois valable que dans la mesure où les données censitaires sont chronologiquement comparables. Toute modification apportée aux questions, concepts et procédures risque de réduire la comparabilité des données.

D'autre part, des erreurs peuvent se répercuter sur l'utilisation des données. On admet couramment l'inévitabilité de la présence, dans les données, de quelques cas de non-réponse ainsi que d'erreurs d'observation, de réponse, de traitement et autres anomalies. L'immense superficie du Canada et la mobilité de ses habitants, de même que la diversité des caractéristiques sociales et culturelles de sa population, risquent d'engendrer dans les recensements, des erreurs qui nécessitent une étude spéciale, car on utilise de plus en plus les données censitaires à l'égard des petites régions géographiques et des sous-groupes de population.

Pour favoriser l'utilisation la plus efficace possible des données, il faut que le produit du recensement englobe des renseignements sur les concepts et méthodes qui sous-tendent chronologiquement les données, ainsi que des renseignements sur la qualité de ces censitaires.

Le présent document a pour objet de décrire le programme d'évaluation que le Bureau a entrepris de réaliser au sujet du recensement national de 1981. Ce programme visait à fournir aux utilisateurs des renseignements fondamentaux sur la comparabilité et la qualité des données censitaires.

2. Stratégie d'évaluation du recensement de 1981.

L'évaluation des données du recensement de 1981 a été planifiée et effectuée d'une manière compatible avec les pratiques qui s'appliquaient aux recensements antérieurs, mais on s'est en même temps efforcé d'élargir et de reformuler ces pratiques afin de mieux réaliser l'objectif que nous poursuivons qui est de fournir à l'utilisateur un meilleur service. À cette fin, on a planifié l'évaluation en se fondant sur les cinq principes directeurs suivants:

(i) les chiffres de population et autres données censitaires doivent être analysés du point de vue de la qualité et de la comparabilité historique, avant que le Bureau les publie:

(ii) les anomalies et les modifications de concepts ou de procédures importantes doivent être indiquées au moment de la publication et faire l'objet d'une documentation explicative intégrée aux données diffusées;

(iii) autant que possible, il faut produire des estimations d'erreurs dans le cas des sources importantes d'erreurs possibles;

(iv) il importe que ces estimations aient un sens pour l'utilisateur et qu'elles fournissent une indication des conséquences potentielles de l'erreur sur l'utilisation des données en cause;

(v) une évaluation de la qualité des données censitaires et de l'impact des méthodes et procédures du recensement sur la qualité et la comparabilité de ces données doit aussi faire l'objet d'une documentation diffusée en tant que partie intégrante du programme des publications courantes du recensement.

Un problème de délai s'est posé dans le cas des deux premiers principes directeurs. L'évaluation de données fondée sur des études statistiques qui permettent de faire des estimations d'erreurs est un travail de longue haleine. Empêcher la diffusion des données censitaires jusqu'à ce que ces mesures estimatives soient connues n'aurait pas été une décision acceptable aux yeux des utilisateurs, et cela aurait sérieusement nui à la pertinence et à la crédibilité des données censitaires comme indicateurs courants des caractéristiques et des répartitions de la population.

En ce qui regarde les troisièmes et quatrième principes directeurs, il existait, du point de vue des coûts, du délai et de la faisabilité, des limitations strictes de l'étendue de l'évaluation possible. Des contraintes analogues se faisaient sentir sur le niveau de détail géographique et de sous-groupement de la population pour lequel on pouvait raisonnablement envisager de faire une évaluation.

Afin de vaincre ou d'atténuer ces constraintes, nous avons abordé le travail d'évaluation sous deux angles, à savoir la certification et la mesure de la qualité des données. La certification visait à nous permettre de respecter les deux premiers principes directeurs, c'est-à-dire la nécessité d'établir, avant la publication et la diffusion des données, une connaissance fondamentale de la qualité de ces dernières et de l'impact des changements de procédures ou de concepts sur la comparabilité chronologique. Comme résultat de l'effort déployé pour fournir une première indication sur la qualité des données, le processus de certification a fait ressortir certains problèmes à propos desquels une évaluation moins rapide mais plus détaillée et plus rigoureuse pouvait être réalisée dans le cadre du processus de mesure de la qualité des données.

La mesure de la qualité des données visait à nous permettre de respecter les troisième et quatrième principes directeurs, c'est-à-dire fournir des mesures fondamentales de la qualité des données et produire des informations quantitatives touchant l'impact des opérations du recensement sur les données résultantes. Il s'agissait notamment de déterminer les sous-populations comportant une erreur appréciable ou possible, de mesurer cette erreur, et de faire une étude détaillée des problèmes constatés au cours du processus de certification.

Pour respecter le dernier principe directeur, nous avons inclus dans le programme officiel de produits et services du recensement, des publications consacrées à la documentation et à l'étude de la qualité des données.

3. Certification

Le processus de certification avait pour objet de cerner les problèmes de qualité qui nécessitaient une rectification avant la diffusion des données en cause, la communication d'un avis approprié aux utilisateurs lors de la diffusion, une étude plus poussée des problèmes constatés ou une combinaison quelconque de ces démarches. Dans le cas des principales répartitions géographiques, nous y sommes parvenus au moyen d'une macro-évaluation des dénombrements des univers (c'est-à-dire dénombrements de personnes, de logements, de familles, etc.) ou proportions, des distributions selon des variables uniques ou multiples, et de mesures découlant de ces répartitions. Ces évaluations ont été facilitées principalement par des comparaisons avec d'autres sources de données et avec des données de recensements antérieurs, soit par rapprochement direct des répartitions, soit par recours à des projections expressément établies à cette fin.

3.1 Étendue de la certification

Comme il y avait plus d'une centaine d'univers et de variables à certifier, et que les données sur ces univers et variables seraient produites à l'égard de différents niveaux de détail géographique, les délais d'exécution prévus nous ont permis de vérifier uniquement les principaux agrégats. Pour la plupart des variables, la certification a été faite à l'échelon national et provincial, mais seulement infraprovincial dans le cas où la nature des vérifications provinciales de la variable était sans signification (par ex., lieu de travail), ou bien lorsque la variable revêtait une importance spéciale, ou lorsqu'il existait une exigence juridique nécessitant une certification infraprovinciale. Ce dernier cas englobait les dénombrements de population (vérifiés au niveau du secteur de recensement - 3,000 à 10,000 personnes) et la variable "langue maternelle", vérifiée au niveau de la circonscription électorale fédérale (c'est-à-dire chaque zone représentée par un député à la Chambre des communes).

Dans le cas de toutes les variables mises à la disposition des utilisateurs sous forme de tableaux publiés ou sur demande spéciale, les répartitions à variable unique ont été certifiées. On a aussi certifié les répartitions à variables multiples dans le cas où il s'agissait de variables fréquemment associées les unes aux autres, ou pour lesquelles il existait une certaine corrélation prévisible.

3.2 Méthode de certification

Ce sont des groupes de spécialistes qui ont procédé à la certification: chaque groupe s'est occupé d'un ou de plusieurs ensembles des univers ou variables corrélés. Avant que les données du recensement soient communiquées à ces groupes, on a choisi d'autres sources de données, ainsi qu'une stratégie de vérification et d'acceptation ou de rejet.

Diverses sources de rechange ont été utilisées. Exclusion faite des recensements antérieurs, ces sources comprenaient des enquêtes de Statistique Canada, des données administratives ou d'enquêtes provenant des ministères ou organismes fédéraux, provinciaux et municipaux, ainsi que des données tirées de différentes publications à caractère spécialisé ou technique.

Les recensements de 1971 et de 1976 sont des exemples typiques des sources de rechange utilisées. Pour les variables "logement" et "ménages", on s'est servi de données remontant au recensement de 1961. Parfois, les données de référence sur les familles remontaient jusqu'au recensement de 1941.

On a mis en oeuvre deux stratégies générales de vérification. La première consistait en une comparaison entre les données du recensement de 1981 et celles de la source de rechange. Des seuils de tolérance ont été prédéterminés pour ce qui concerne le degré d'écart acceptable entre les répartitions en cause et les répartitions comparables tirées des sources de rechange. On a établi ces seuils de tolérance en tenant compte de la comparabilité des concepts et des dates réelles, ainsi que des limitations de l'univers, de la méthodologie, de la qualité des données et d'autres aspects de la source de rechange. Par exemple, la vérification des variables âge/sexe/état matrimonial a fait appel à des estimations démographiques indépendantes comme sources de rechange. En ce qui regarde les données censitaires, on a établi les seuils d'écart tolérable à plus ou moins 5% de l'estimation démographique comparable pour tous les effectifs de 100,000 ou plus, et de plus ou moins 10% pour les effectifs de 10,000 à 100,000. Les effectifs plus petits ont été examinés lorsque le besoin s'en faisait sentir.

Quand un effectif ou toute autre mesure dépassait le seuil de tolérance, il fallait fournir une explication valable de l'écart. Dans quelques cas, l'explication était facile à trouver dans d'autres répartitions des données censitaires (par exemple, en raison de l'ampleur exceptionnelle de la migration interprovinciale, certaines cases âge-sexe au sein d'une province donnée étaient supérieures aux valeurs projetées). D'autres fois, des recherches subséquentes ont révélé que les écarts étaient attribuables à une erreur dans les données du recensement, tandis que dans d'autres cas l'écart était attribuable à une erreur dans les données de la source de rechange.

On a recouru à la seconde stratégie de vérification lorsqu'il n'existait pas de précédent censitaire pour baser la projection, ou lorsque la projection basée sur un recensement antérieur n'était pas entièrement significative, ou encore si les sources de rechange externes n'étaient pas entièrement comparables ou quand il n'en existait pas. Pour ces variables, il a fallu asseoir les répartitions sur une base rationnelle. La vérification interprétative s'est effectuée d'après des critères logiques appliqués aux données. Par exemple, en ce qui concerne la question censitaire posée pour la première fois au sujet de la propriété du logement, on ne s'attendait pas à un genre particulier de réponses, sauf pour certains genre de logements situés en milieu urbain. L'obtention relativement fréquente d'une réponse particulière dans d'autres zones ou à l'égard d'autres genres de logements aurait indiqué une erreur probable.

La stratégie de vérification par comparaison des données a constitué 80% de la certification. Pour le reste, on a recouru à la technique interprétative.

Au cours du processus de certification, toutes les variables ont fait l'objet de comparaisons entre les répartitions relatives aux données imputées et contrôlées et les données non imputées et non contrôlées. Ainsi, on peut dire que, dans une mesure restreinte, nous avons fait une évaluation séparée des méthodes et procédures du recensement.

4. Mesure de la qualité des données

La mesure de la qualité des données a consisté en un examen de l'erreur et (ou) du changement introduit(s) dans le recensement par cinq sources possibles d'erreur - observation, réponse, traitement (autre que le contrôle et l'imputation), contrôle et imputation, échantillonnage et pondération. Ces composantes de l'examen visaient à englober tous les aspects du travail de collecte et de traitement des données censitaires.

À l'égard de chacune de ces sources, on a fait des études en vue de fournir des estimations de l'erreur ou de décrire quantitativement l'impact des opérations ou des procédés sur l'information censitaire. Pour choisir, dans l'ensemble des variables, celles qu'il convenait d'étudier, nous nous sommes fondés sur une évaluation de l'importance, relativement aux utilisations des données censitaires, de l'erreur possible qu'il convenait d'examiner et du rapport coût-efficacité du choix.

4.1 Nature des opérations

Les mesures de la qualité des données ont été reproduites au moyen de quatre types généraux d'étude:

(i) Contrôles du dénombrement: ces contrôles étaient en fait des études statistiques basées sur un échantillon et ils comportaient un certain degré de nouveaux dénombrements. Ces contrôles ont été effectués surtout dans le cadre des études de l'observation.

(ii) Opération parallèles: ces opérations ont consisté en des études réalisées de façon indépendante et qui portaient sur un échantillon de données censitaires, parallèlement à une opération du recensement. La comparaison des résultats du recensement avec ceux de l'opération portant sur l'échantillon a permis d'estimer l'erreur introduite. On a fait ce genre d'études pour évaluer la qualité du travail de traitement.

(iii) Comparaisons avec les autres sources de données: ces comparaisons faisaient appel à des micro-appariements, basés sur un échantillon, des dossiers censitaires avec ceux de la source de rechange. On a effectué ce genre d'études pour évaluer la réponse ainsi que le contrôle et l'imputation.

(iv) Analyse des données du recensement: cette analyse a comporté une certaine forme de calcul qui faisait appel à des données censitaires définitives et (ou) intermédiaires (non contrôlées). On a fait ce genre d'études pour déterminer les taux de réponse, ainsi que pour établir des estimations de l'erreur d'échantillonnage et pour évaluer la qualité du contrôle et de l'imputation.

4.2 La composante "observation"

L'objet premier du recensement est de fournir des chiffres de population selon divers niveaux géographiques. Les erreurs d'observation revêtent donc une importance particulière, de sorte que l'on a consacré une partie assez considérable des ressources évaluatives à l'estimation de ces erreurs.

Il se produit une erreur d'observation chaque fois qu'une personne, un ménage ou une logement échappe au dénombrement (sous-dénombrement), ou lorsque l'entité est comptée plus d'une fois (surdénombrement) ou encore lorsqu'une unité non comprise dans l'univers censitaire est prise en compte (surdénombrement). En général, on s'attend à ce que les cas de surdénombrement soient plutôt rares comparativement aux cas de sous-dénombrement et le surdénombrement est plus difficile à évaluer. La composante "observation" du programme de mesure de la qualité des données a donc été orientée principalement, mais non exclusivement, vers l'estimation de l'erreur attribuable au sous-dénombrement.

On a évalué quatre éléments de l'erreur d'observation:

(i) le sous-dénombrement de la population et des ménages privés, qui a été examiné au moyen de la contre-vérification des dossiers;

(ii) le sous-dénombrement de la population de personnes temporairement absentes de leur lieu de résidence habituelle, qui a été examiné dans le cadre de l'étude sur les résidents temporaires;

(iii) le sous-dénombrement de la population et des ménages composés de personnes dont le lieu de résidence habituelle, avait été, par inadvertance, classé comme vacant; cet élément a été étudié dans le cadre de la vérification des logements vacants;

et (iv) le surdénombrement des logements attribuable au fait que des locaux d'habitation non compris dans le stock de logements (parce qu'ils étaient en construction ou en réfection, ou bien qu'il s'agissait de logements saisonniers, etc.) avaient été classés à tort comme logements vacants.

4.2.1 La contre-vérification des dossiers

La contre-vérification des dossiers (CVD) s'effectue dans le cadre du recensement national depuis 1966. La CVD avait pour but précis (i) de fournir des estimations du sous-dénombrement de la population au niveau provincial; et (ii) de fournir des estimations du sous-dénombrement des ménages privés et de la population dans divers sous-groupes des univers de la population et du logement.

La méthode de CVD comportait deux éléments-clés: (i) l'élaboration d'une liste indépendante de toutes les personnes qui, croyait-on, auraient dû être dénombrées dans le recensement; et (ii) les opérations de dépistage et de suivi, que l'on a effectuées pour repérer sur la liste établie les cas de décès et d'émigration qui se sont produits au cours de la période intercensitaire, et pour s'assurer que l'on connaissait, le jour même du recensement, le lieu habituel de résidence des personnes qui auraient dû être dénombrées.

En principe, ces deux éléments permettaient de faire une comparaison avec les données du recensement, pour déterminer si les personnes en cause auraient dû être dénombrées. Après avoir établi la liste, nous avons procédé au travail de dépistage, de suivi et d'appariement à l'égard d'un certain échantillon de personnes.

La liste indépendante comportait quatre éléments constitutifs ou colonnes:

(i) toutes les personnes dénombrées lors du recensement précédent (1976);

(ii) toutes les naissances intercensitaires - d'après les dossiers de la statistique de l'état civil:

(iii) tous les immigrants intercensitaires - d'après les dossiers de l'immigration;

et (iv) toutes les personnes omises lors du recensement précédent (1976) - d'après la CVD de 1976.

On a tiré un échantillon de personnes de toutes les colonnes sauf la quatrième, étant donné que cette colonne était déjà représentée par un échantillon de cas résultants de la CVD de 1976 et qui avaient été classés comme oubliés dans le recensement de cette année-là. L'échantillon total comptait environ 36,500 personnes.

Une fois l'échantillon sélectionné, on s'est livré à de vastes travaux pour dépister les personnes choisies et en assurer le suivi. Dans le cas de presque toutes ces personnes, on a cherché dans les documents censitaires (par exemple, du recensement antérieur) une adresse plausible. Si la personne choisie n'était pas dénombrée à cette adresse, on tentait de communiquer avec elle, par des entrevues téléphoniques ou par consultations d'annuaires de ville ou de différents dossiers administratifs gouvernementaux. On a recherché l'adresse du jour du recensement de celles qui ont été contactées. Si une personne choisie n'a pas été denombrée à l'adresse donnée, on procédait à une interview de suivi pour obtenir des informations sur d'autres adresses possibles et pour recueillir certaines données censitaires concernant cette personne et son ménage. On utilisait ensuite ces informations pour confirmer si oui ou non la personne en cause avait été oubliée lors du recensement, et dans l'affirmative, l'information servait d'apport aux estimations du sous-dénombrement.

Par suite de ces opérations, les cas de l'échantillon ont été classés dans l'un ou l'autre des cinq résultats fondamentaux suivants: (i) personnes dénombrées lors du recensement, (ii) oubliées lors du recensement, (iii) décédées avant le jour du recensement, (iv) émigrées avant le jour du recensement ou vivant à l'étranger ce jour-là, et (v) personnes choisies mais non dépistées. La catégorie des personnes non dépistées correspondait en fait à des "non-réponses" et représentait les cas pour lesquels l'opération de dépistage n'avait rien donné. Parmi toutes les personnes choisies, 3.4% ont été classées dans cette catégorie. On a supprimé ces cas de l'estimation subséquente, au moyen des procédures de rajustment de la pondération.

4.2.2 Étude des résidents temporaires

Le recensement du Canada porte sur la population "légale" et non sur la population "de fait". Ainsi, les personnes temporairement absentes de leur lieu de résidence habituelle doivent encore être dénombrées à ce lieu. Toutes les personnes qui, le jour du recensement national, faisaient partie de l'univers censitaire (à l'exclusion des résidents étrangers) et qui se trouvaient temporairement dans une résidence quelconque (par exemple, hôtel, hôpital ou prison) ont été dénombrées à cette résidence au moyen d'une formule spéciale du recensement, apprlée formule 3.

Après l'opération de collecte, on a choisi un échantillon constitué de 2% des formules 3 remplies (environ 10,000), puis on a entrepris des travaux de dépistage pour trouver dans les questionnaires normaux du recensement le lieu de résidence habituelle de chaque personne choisie. Ces travaux ont été facilités par l'information que la formule 3 demandait au sujet de l'adresse. Il s'agissait ensuite de déterminer, à l'aide de l'information recueillie (nom, âge, sexe et état matrimonial) si la personne en cause avait été dénombrée à son lieu de résidence habituelle.

Le Bureau a produit des estimations du nombre de personnes oubliées - selon le groupe d'âge, le sexe et l'état matrimonial - par grande région géographique comprise à l'intérieur des provinces. On a utilisé ces estimations pour rajuster ou "rectifier" les chiffres de population dans ce cas particulier de sous-dénombrement. Cette rectification s'est effectuée au moyen de rajustment de la pondération faisant appel à des enregistrements donneurs choisis au hasard et présentant certaines caractéristiques (âge, sexe, état matrimonial) à l'intérieur de régions géographiques déterminées. Par ce processus, le poids de l'enregistrement était augmenté de 1, si bien qu'il représentait alors sa propre entité et celle de la personne oubliée. Ainsi, on a pu ajouter près de 118,000 personnes aux chiffres initiaux de population, soit environ 0.5% du chiffre définitif de population et 23% de toutes les personnes dénombrées au moyen de la formule 3 à leur lieu de résidence temporaire.

C'est parce que les personnes dénombrées comme résidents temporaires formaient un groupe aisément identifiable et connu pour être sujet à un taux élevé de sous-dénombrement que l'on a entrepris cette étude et procédé par la suite au rajustment nécessaire.

4.2.3 Contrôle des logements vacants

Les logements inscrits comme vacants formaient, dans l'univers du recensement, un autre groupe manifestement sujet à un taux assez élevé d'erreur d'observation. En l'occurrence, il convenait d'examiner deux aspects de l'erreur d'observation. Tout d'abord, les logements inscrits comme vacants étaient peut-être, en fait, occupés le jour du recensement. Le cas échéant, il pouvait s'ensuivre un sous-dénombrement de personnes, de ménages et de logements occupés, et un surdénombrement des logements vacants. Une telle anomalie de classement pouvait résulter de ce que le ménage avait déménagé à une date rapprochée de la journée du recensement ou qu'il était absent à ce moment-là (par exemple, il pouvait être en vacances ou en séjour dans une résidence d'été). Deuxièmement, les constructions dénombrées comme logements vacants ne se prêtaient peut-être pas à une occupation permanente le jour du recensement, de sorte qu'il n'aurait pas fallu les classer comme logements. Le cas échéant, il en résultait un surdénombrement aussi bien des logement

vacants que du stock total de logements.

Afin d'estimer ces éléments de l'erreur d'observation, on a effectué, peu après l'opération de collecte des données du recensement, un nouveau dénombrement par sondage des logements ou constructions initialement dénombrés comme vacants. Aux fins de ce sondage, on a choisi le travail d'environ 1,250 agents recenseurs, ce qui représentait quelque 14,000 "logements vacants".

Le nouveau dénombrement s'est effectué sous forme d'interviews intensives avec les occupants de l'heure (lorsque le logement était alors effectivement occupé), les concierges et les voisins. À la lumière des résultats de ces entrevues, on a pu établir si la construction en cause était effectivement occupée le jour du recensement. Dans l'affirmative, la recherce des adresses obtenues a permis de déterminer si les occupants avaient été dénombrés ailleurs. On a également pu déterminer le nombre de personnes composant le ménage, et le genre de logement (par exemple, immeuble d'appartements).

Dans le cas où la construction visée était vacante le jour du recensement, on a établi s'il s'agissait ou non d'un logement. Pour être considérée comme logement, il fallait que le jour du recensement elle put être habitée à longueur d'année. Autrement dit, il fallait qu'elle comportât certaines installations essentielles. Par exemple, une maison en construction ou en renovation et un chalet d'été non chauffé ne devaient pas être considérés comme des logements, de sorte qu'il ne fallait pas les dénombrer.

L'étude faite à cet égard a permis d'établir deux ensembles primaires d'estimations de l'erreur d'observation. Le premier ensemble se composait d'estimations du nombre de ménages oubliés, selon la taille du ménage, le genre de logement et les régions urbaines-rurales des diverses provinces. (L'estimation du nombre de ménages oubliés représentait aussi un volume correspondant de surdénombrement des logements vacants.) On a utilisé ces estimations pour rajuster ou "rectifier" les chiffres censitaires de population, de logements occupés (ménage) et de logements vacants, dans ce cas particulier de sous-dénombrement et de surdénombrement. La rectification des chiffres de population et de logements occupés a été effectuée au moyen de processus de rajustement de la pondération analogue. grosso modo, au processus utilisé pour rectifier le nombre de résidents temporaires. Dans ce cas, on a recouru à des enregistrements donneurs qui étaient des enregistrements de ménages choisis en fonction de la taille du ménage et du genre de logement à l'intérieur de régions géographiques déterminées. Parallèlement, on a retiré de la base de données un nombre correspondant de logements vacants. Par cette méthode, on a ajouté environ 44,500 ménages et 82,000 personnes au chiffres initiaux de population et de ménage. Cela représentait quelque 0.3% du chiffre définitif de population et correspondait, dans le cas des constructions inscrites comme vacantes, à un taux d'erreur de classement de 10%.

La seconde ensemble se composait de diverses estimations du nombre de constructions classées à tort comme logement. Ces estimations n'ont pas servi à des fins de rectification, mais on en a fait mention dans une publication de données censitaires pour alerter l'utilisateur au sujet des cas possibles de surdénombrement dans les chiffres du stock total de logements et dans le nombre de logements vacants. Après rectification des chiffres en cause, on a estimé que le taux de surdénombrement des logements vacants atteignait presque 20%. La moitié environ de ce pourcentage était attribuable à un dénombrement

fautif des logements en construction ou en renovation.

4.2.4 Erreur nette de sous-dénombrement

La CVD par sondage des données censitaires a simplement révélé que les chiffres du recensement étaient complets avant les rectifications apportées dans le cas des résidents temporaires et des logements inscrits à tort comme vacants. Il a donc fallu rajuster les estimations de la CVD pour justifier les ajouts de statistiques à la base de données. Ce rajustment a révélé que le taux net du sous-dénombrement de la population en 1981 s'établissait à 2%. C'est là un niveau de sous-dénombrement analogue à celui des trois recensements précédents.

4.3 La composante "réponse"

On a fait trois études générales pour évaluer la qualité des réponses - soit une étude du taux de réponse, une étude du contenu basée sur l'échantillon de la CVD, et l'investigation de divers problèmes.

4.3.1 Étude du taux de réponse

On a établi les taux de réponse nationaux pour chaque variable prise en compte. Même si ces taux de réponse ne constituaient que des indicateurs généraux du niveau de la qualité, ils n'en révélaient pas moins la mesure dans laquelle les données du recensement reposaient sur les réponses du public, et, jusqu'à un certain point, le volume d'imputation requis.

À l'égard d'un échantillon de dossiers censitaires, on a calculé des taux plus détaillés de réponse à variable unique et à variables multiples. Ce calcul avait pour objet de déterminer les groupes qui, au sein de la population, présentaient des taux élevés de non-réponse. Un taux élevé de non-réponse indiquait la présence possible d'un biais dans les données concernant le groupe observé. De même, pour la majorité des groupes en cause, un taux minime de non-réponse indiquait une faible possibilité de biais attribuable aux cas de non-réponse.

4.3.2 Contenu de l'étude sur la contre-vérification des dossiers

Cette étude visait tous les dossiers choisis de personnes que la CVD avait classées comme "dénombrées" en 1981. Relativement à ces dossiers, on a fait des comparaisons entre les réponses fournies par la personne au recensement de 1981 et celles obtenues lors du recensement 1976 (ou tirées de dossiers administratifs, dans le cas des personnes non dénombrées en 1976). Les comparaisons ont porté sur les variables qui ne sont pas censées se modifier avec le temps (par exemple, date de naissance, sexe, langue maternelle) et sur les variables qui ne peuvent se modifier que dans une seule direction (par exemple, plus haut niveau de scolarité atteint). Comme mesures primaires, on a produit des estimations du biais relatif et des taux bruts de différence.

4.3.3 L'analyse des problèmes

En raison des problèmes mis en évidence par le processus de certification, on a entrepris certaines études particulières. De façon générale, ces études ont eu pour résultat une définition explicite du problème, sans nécessairement aboutir à des estimations du degré précis d'erreur.

4.4 La composante "traitement"

On a fait deux études en vue d'évaluer les travaux de codage, de saisie de données et de contrôle manuel.

4.4.1 Étude du biais attribuable au codage

Des travaux de codage s'imposaient pour transformer diverses incriptions littérales (par exemple, branche d'activité, profession) en codes numériques pouvant ensuite être saisis. On a donc fait une étude du biais introduit par les codeurs.

Pour faciliter l'estimation du biais attribuable au codage, on a comparé un échantillon des codages effectués par les codeurs avec ceux effectués par des codeurs "experts". Cela a été fait dossier par dossier pour chaque variable. En produisant les estimés, nous avons supposé l'exactitude des codeurs experts. Dans une mesure indéterminée, la fiabilité des résultats était limités dans cette hypothèse, du fait que seul un échantillon de dossiers de personnes pouvait servir à l'étude, et aussi parce que les cas d'inscriptions en toutes lettres sont rates en ce qui concerne certaines variables et que les codes individuels ne sont pas non plus fréquents.

4.4.2 Évaluation de la saisie des données

Les données censitaires ont été enregistrées par clavier. Comme cette méthode de saisie n'avait pas été utilisée lors du recensement précédent, il était donc important de montrer que cette façon de procéder ne constituait pas une source appéciable d'erreurs.

Aux fins de l'étude évaluative faite à ce sujet, on a procédé, en dehors du flux principal des données censitaires, à une nouvelle saisie des réponses figurant sur les questionnaires du recensement d'un échantillon qui représentait 33,000 personnes. Les données saisies de nouveau ont ensuite été traitées dans le reste du système de saisie, qui prévoyait des contrôles manuels fondamentaux et des contrôles du nombre de personnes dénombrées. Par suite de ce traitement, des spécialistes ont apporté les modifications que nécessitaient les données de l'échantillon.

Après les avoir traités de nouveau, on a comparé les dossiers de l'échantillon aux dossiers correspondants saisis lors du traitement principal des données censitaires. Tous les écarts entre les deux versions de données ont été signalés aux experts, et les discordances attribuables à une erreur commise dans le traitement principal a fait l'objet d'un bit indicateur. On a ensuite établi des estimations de l'erreur de traitement, par variable, au moyen de drapeaux indicateurs d'erreur.

4.5 La composante "contrôle et imputation"

On a fait trois études générales pour évaluer l'impact de contrôle et de l'imputation sur les données censitaires. Ces études ont été mises en parallèle avec celles portant sur l'évaluation du taux de réponse.

4.5.1 Relations entre les données

L'étude sur les relations entre les données a fourni des mesures de l'ampleur des imputations - tant à l'égard des lacunes de contrôle que des cas de non-réponse. Même si ces mesures ne constituaient pas des estimations directes de l'erreur, elles ont fourni une forte indication de la qualité des données. Vu que le processus de contrôle et d'imputation vise les erreurs détectables (y compris les cas de non-réponse) dans tout ensemble fixe de procédure de traitement, plus la proportion de changement ou d'imputation est faible, plus les données devraient être fiables. Jusqu'à un certain point, cela doit évidemment conduire à une évaluation des règles même du traitement.

On a établi des estimations du taux d'imputation pour chaque variable, en fonction de chaque valeur finale. Sur une base de variables multiples, on a aussi calculé des estimations analogues en faisant appel à des combinaisons de valeurs imputées et de valeurs non imputées. On s'est ensuite servi de ces estimations pour identifier les groupes de population à l'égard desquels il s'était produit un taux élevé d'imputation, et pour déterminer quelle était la tendance à l'imputation dans le cas de certaines valeurs. En partie, cela constituait une évaluation des règles suivis en matière de contrôle et d'imputation.

Cet effort d'evaluation a été guidé dans une certaine mesure par le travail de certification, qui a permis de déterminer les secteurs où le contrôle et l'imputation avaient eu un impact sensible sur les données, et par les résultats de l'étude sur les taux de réponse.

4.5.2 Appariement des résultats de la contre-vérification des dossiers

Une fois le processus d'appariement terminé aux fins de l'étude sur le contenu de la CVD, on a pu utiliser les mêmes dossiers appariés pour évaluer l'impact du contrôle et de l'imputation. Essentiellement, on a recouru pour cela à la même méthode que pour l'étude du contenu de la CVD, sauf qu'au lieu d'apparier les réponses initiales du recensement, on a comparé les réponses de la CVD (soit du recensement antérieur ou des dossiers administratifs, ou celles recueillies dans le cours même de la CVD) avec les valeurs imputées, le cas échéant.

On a établi des estimations distinctes pour le travail d'imputation qu'exigent, d'une part, la résolution des discordances révélées par le contrôle, et, d'autre part, les cas de non-réponse. Les estimations produites pour l'examen de discordances ont été évaluées en regard des résultats correspondants du contenu de la CVD.

4.5.3 Analyse des problèmes

Pour examiner les problèmes réels de contrôle et d'imputation, on a pro-

cédé de la même manière que pour l'étude des problèmes de réponse suspectés. En général, on cernait le problème au cours de la certification ou à la lumière d'une autre étude, ce qui donnait lieu à un examen détaillé de la procédure de contrôle et d'imputation, compte tenu des circonstances de la réponse et du travail de collecte.

4.6 La composante "échantillonnage et pondération"

Il était entendu que les données fondamentales du recensement de 1981 seraient recueillies à l'égard de la totalité de la population, mais un volume considérable d'informations ne devait être obtenu que d'un échantillon aléatoire de ménages (un sur cinq). Il fallait donc évaluer l'impact de l'échantillonnage et du processus subséquent de pondération sur la qualité des données recueillies. Aux fins de cette évaluation, on a procédé à trois études primaires.

4.6.1 Contrôle du biais de l'échantillonnage

Les exigences du plan de l'échantillon nécessaire au recensement et les exigences des procédures correspondantes de sélection de l'échantillon, que le recenseur devait suivre, étaient censées garantir que l'échantillon choisi ne présenterait pas de biais. On a donc effectué un contrôle du biais possible de l'échantillon pour déterminer dans quelle mesure l'échantillon avait été choisi en conformité avec le plan voulu.

Aux fins de cette évaluation, on a comparé, pour ce qui est de tous les secteurs du dénombrement (c'est-à-dire chaque unité géographique confiée à un agent recenseur), la répartition des caractéristiques des ménages de l'échantillon choisi avec les caractéristiques de tous les ménages du SD (en recourant à toutes les données recueillies auprès de tous les ménages). On a soumis ces comparaisons à des essais statistiques pour déterminer si les différences, parmi tous les SD de certaines régions géographiques, s'écartaient sensiblement des chiffres auxquels ont pouvait s'attendre dans le cas d'un échantillon aléatoire simple de ménages.

4.6.2 Contrôle de la concordance échantillon/population

On a appliqué la procédure de pondération censitaire à des groupes de SD contigus appelés zones de pondération (ZP). La ZP typique comptait une population de 3,000 à 7,000 personnes. La méthode de pondération faisait appel, dans le cadre d'une procédure d'estimation du ratio de classement, à diverses données sur la population et le logement recueillies à l'égard de tous les ménages. En l'occurrence, il s'agissait, en quelque sorte, d'une procédure d'estimation à deux dimensions, dont chacune était appliquée de façon alternative ou itérative. À l'égard des caractéristiques utilisées dans ce processus de rajustment, les estimations basées sur les données du sondage, au niveau de la ZP ou à un niveau supérieur, devaient se rapprocher étroitement des chiffres correspondants que l'on aurait obtenus de la totalité correspondante des personnes ou des ménages. En ce qui concerne les données recueillies auprès de l'univers total, cette procédure devait normalement révéler un fort degré de concordance entre les totalisations basées sur les données de l'univers complet et les totalisations basées sur les données de l'échantillon.

On a procédé au contrôle de la concordance échantillon/population pour mesurer l'étendue de cette concordance et, en particulier, pour déterminer si la concordance au niveau des ZP avait été obtenue au détriment de la concordance aux niveaux inférieurs de pondération. Pour réaliser cette évaluation, on a comparé, dans le cas de diverses répartitions géographiques allant jusqu'au niveau du SD, les chiffres provenant de la totalité de l'univers pour une gamme déterminée de variables concernant cet univers aux données estimatives correspondantes tirées de l'échantillon. Tout comme dans l'étude du biais de l'échantillonnage, on a soumis ces comparaisons à des tests statistiques.

4.6.3 Erreur d'échantillonnage

Il était essentiel de mesurer l'erreur introduire dans les données censitaires par l'échantillonnage. Pour la plupart des variables du recensement, on a établi des estimations de la variance (intégrant la variance d'échantillonnage et quelques composantes de la variance attribuable à la réponse et au traitement). Les estimés ont été basés sur un échantillon de 10% des ZP, et ont été établis en utilisant un estimateur basé sur les termes du développement d'une série de Taylor, qui tient compte de l'effet de la procédure de pondération. Des ensembles distincts d'estimations ont été établis, par province.

5. Publication des renseignements sur la qualité des données

Les renseignements sur la qualité des données ont été (ou seront) diffusés de diverses manières. Toutes les données rendues publiques dans le cadre du programme établi de produits du recensement étaient accompagnées d'un texte normalisé sur la qualité des données elles-mêmes. Ce texte comprenait un exposé des sources possibles d'erreur et signalait à l'utilisateur que les données pouvaient être plus ou moins fautives, notamment dans le cas des petites valeurs. Parfois, le texte comprenait aussi des estimations de l'erreur au niveau national (par exemple, les taux estimatifs du sous-dénombrement de la population et des ménages privés, et les taux de réponse).

Le texte d'accompagnement des données recueillies par sondage comprenait un exposé de l'erreur potentielle attribuable à l'échantillonnage, ainsi qu'un tableau des erreurs d'échantillonnage (répartis par taille de valeur) assorti d'indications sur l'emploi de ces mesures. En outre, si les données diffusées englobaient des tranches de revenu moyen, les tableaux de données comportaient aussi des estimations correspondantes de l'erreur type des estimations du revenu moyen.

Bien entendu, ces informations s'ajoutaient aux divers genres de notes mentionnées ci-dessus, et aux définitions fondamentales des variables ainsi qu'à la terminologie du recensement. Cependant, le texte informatif variait en fonction des données comprises dans l'ensemble de totalisations en cause.

Pour ce qui est des données fournies à tel ou tel utilisateur sur demande, un texte explicatif, y compris les mises en garde appropriées, étaient envoyées avec les totalisations désirées. En outre, au moment de la demande, on examinait à fond avec l'utilisateur la qualité des données recueillies, ainsi que les questions de comparabilité, de définition, etc.

Tous les textes accompagnant les totalisations censitaires, aussi bien

que tous les autres bulletins pertinents, renvoyaient l'utilisateur à deux publications consacrées à l'examen critique de la qualité des données du recensement. Ces deux publications représentent la forme la plus complète d'étude de la qualité des données censitaires, de leur comparabilité historique et de l'impact des méthodes du recensement de 1981 sur la qualité et la comparabilité de ces données. La première publication porte sur les données recueillies auprès de toute la population, tandis que l'autre concerne les données recueillies auprès d'un échantillon de ménages.

Dans ces publications, on décrit brièvement les aspects de la méthodologie censitaire qui influent sur la qualité et la comparabilité des données, et on présente les résultats du travail de mesure de la qualité des données, par source possible d'erreur. Toutefois, la majeure partie de ces deux documents consiste en une évaluation des données censitaires, univers par univers et variable par variable. Cette évaluation se fonde sur une analyse des résultats de la certification, réexaminés à la lumière des résultats de l'effort de mesure de la qualité des données.

6. Efficacité du programme d'évaluation

Présentement, nous poursuivons le travail de planification et de mise au point indispensable à l'évaluation du recensement de 1986. Ce travail repose, entre autres, sur une étude de l'efficacité du programme d'évaluation des données de 1981.

6.1 Certification

En ce qui concerne le recensement de 1981, le travail de certification visait à évaluer la continuité chronologique et la cohérence des données censitaires alors recueillies. Il s'agissait de détecter la présence de tout biais appréciable dans ces données. Pareil biais est détectable uniquement dans les conditions suivantes:

(i) s'il n'existait pas de biais chronologique et systématique, ni de biais commun avec les autres sources de données;

(ii) si les concepts, méthodes et autres éléments propres aux sources de rechange avaient été bien compris;

(iii) si les projections utilisées à des fins de comparaison étaient bien formulées et réellement valides;

(iv) si les limites de tolérance correspondaient, dans une mesure raisonnable, à des écarts autres que des biais par rapport à la source de rechange; et

(v) si le biais était assez considérable pour ne pas être masqué par des variances liées aux données censitaires et aux données provenant de la source de rechange.

La certitude que ces diverses conditions étaient remplies dans le cas de chacune des variables considérées devait être fournie, soit par une preuve raisonnable tirée du travail de certification établissant que les données re-

latives à la variable en cause contenaient un biais appréciable, ou par les résultats des études de mesure de la qualité des données. La certification constituait une approche efficace pour les variables ayant déjà fait l'objet d'études approfondies, et pour les variables qui se révèlent stables à travers le temps ou dont l'évolution est relativement facile à prévoir (par exemple, âge, sexe, et variables du logement). En général, les résultats de la mesure de la qualité des données concernant ces catégories de variables corroboraient les conclusions de la certification. Dans deux cas importants, les résultats de la certification n'ont pas été corroborés. En l'occurrence, il existait dans les données censitaires un biais historique (réponse ou recenseur) qui invalidait les comparaisons de la certification.

La méthode dite de certification s'est révélée moins efficace pour les variables culturelles et d'autre variables susceptibles d'importantes erreurs de réponse, qu'il était difficile d'évaluer. En général, on ne disposait pas pour ces variables, d'autres sources de données valables ou comparables, ou bien l'on ne parvenait pas facilement à prévoir l'évolution temporelle des renseignements en cause. Souvent aussi on n'avait pas de mesures significatives de la qualité des données, qui auraient permis de déterminer l'erreur de réponse liée à ces variables. Dans ces divers cas, les conclusions de la certification demeurent peut-être encore discutables.

Le processus de certification n'écartait donc pas tout risque que nous soyons amenés à rendre public des données entachées d'une erreur importante, mais inconnue (et non suspectée). Toutefois, la certification a considérablement réduit le risque subsistant. D'autre part, le travail d'analyse qui faisait partie de la certification nous a fourni, avant l'étape de la publication, une connaissance des phénomènes (réels ou autres) que les données censitaires décrivaient. Cela a directement contribué à accroître l'efficacité de nos communications avec les utilisateurs.

6.2 Mesure de la qualité des données

Pour la plupart, sinon pour tous les aspects du recensement que nous avons examinés, les études de mesure de la qualité des données ont fourni des résultats significatifs et utiles. Les limitations de cette partie du programme avaient trait à son ampleur ou à son niveau d'application. Trois limitations méritent d'être mentionnées ici. Tout d'abord il existe d'importantes sources possibles d'erreur pour lesquelles on ne dispose pas d'estimations. Par exemple, dans le cas de nombreuses variables, on n'a en main ni estimation du biais attribuable aux réponses mêmes, ni estimation du surdénombrement de la population ou du nombre d'étrangers en situation irrégulière, dont la présence risquait de fausser les chiffres de population.

En second lieu, les estimations d'erreur ne s'appliquent généralement pas au données sur les petites régions. Même si quelques estimés (par exemple, des estimés sur l'erreur d'échantillonnage) semblent applicables pour n'importe quelle taille de groupe, il est toujours possible que l'erreur réelle soit grandement affectée par des circonstances spéciales ou locales. En partie, c'est là un problème que pose l'erreur d'échantillonnage liée aux estimations de la qualité des données.

Troisièmement, et jusqu'à un certain point la limitation dont il s'agit ici se confond avec la précédente, il n'existe que peu d'estimations ou de me-

sures de l'erreur totale ou agrégative. Cette lacune ne posera pas nécessairement de problème sérieux pour ce qui est des cases à valeur élevée, pourvu que l'on puisse déterminer les sources possibles d'erreur (indépendamment du niveau même de l'erreur). Toutefois, la difficulté s'accentuera à mesure que les tailles de groupe diminueront, et elle se fera particulièment aiguë au niveau des données sur les petites régions.

7. Conclusion

Le programme d'évaluation du recensement de 1981 comprenait d'importantes améliorations par rapport aux programmes d'évaluation des recensements antérieurs. La portée et le détail des travaux d'évaluation des données censitaires de 1981 avaient un peu plus d'ampleur. Toutefois, les améliorations les plus sensibles ont résidé dans l'application d'un processus de certification officiel, ainsi que dans l'objet, la nature et la présentation des renseignements destinés aux utilisateurs.

De toute évidence, on peut conclure qu'en vue du recensement de 1986 il faut résoudre les problèmes qui limitaient le programme d'évaluation de 1981. Cependant, il ne faut pas oublier que l'amélioration des travaux d'évaluation de la qualité des données est un processus évolutif qui se déroule dans le cadre de changements révolutionnaires de l'emploi et du traitement des données. Sans doute ne parviendra-t-on pas bientôt à vaincre les servitudes qui restreignaient le programme de 1981, tandis que la demande de données s'intensifiera de plus en plus. Il est donc essentiel que l'on se préoccupe sans cesse de la qualité des données dans la mise au point des questionnaires, des méthodes et des opérations de recensement. Ainsi, la qualité des données censitaires s'accroîtra au maximum, même s'il est impossible de la mesurer avec précision.

POOR

QUALITY

ORIGINAL